Subequivariant Graph Reinforcement Learning in 3D Environments

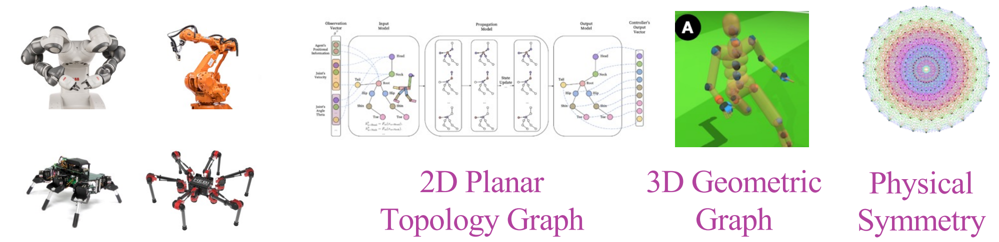
Runfa Chen*, Jiaqi Han*, Fuchun Sun, Wenbing Huang

* Equal contribution

ICML 40 Years International Conference On Machine Learning

Background

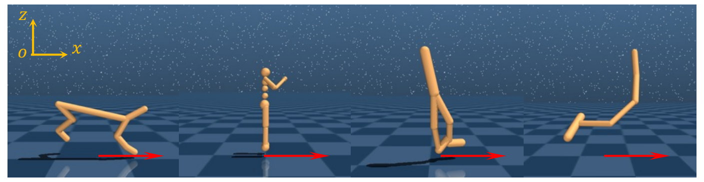
❑ Challenges posed by morphology



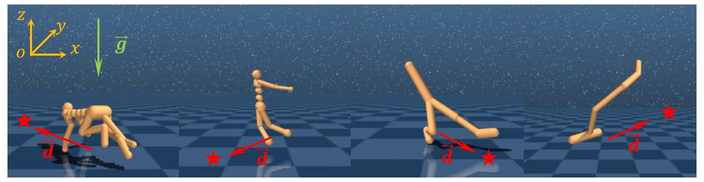
- ➢Each robot has a different morphology.
- ➢A separate policy is trained for each robotics setup. It doesn't generalize.
- ➢Prior Attempts only use topology graph in 2D Planar environments.
- ➢Real Physical World is 3D Geometric Structure and Systems, which contains Physical Symmetry.

Motivation: 3D-SGRL

❑ Illustrative comparison between previous 2D planar setting and our 3D subequivariant formulation.



(a) 2D Planar Locomotion Environments



(b) 3D Subequivariant Locomotion Environments

		2D-Planar	Our 3D-SGRL
State Space	Range	xoz-plane	3D space
	Initial	x^+-axis	Arbitrary direction
	Target	x^+-axis	Arbitrary direction
Action Space	# Actuators	1 per joint	3 per joint
	DoF	1 per joint	3 per joint
Symmetry	External Force	NULL	Gravity $\vec{g}$, Target $\vec{d}$
	Group	$\emptyset$	$\mathrm{O}_{\vec{g}}(3)$

Method: SubEquivariant Transformer (SET)

❑ Equivariance

Definition 2.1 (E(3)-equivariance). Suppose $\vec{\boldsymbol{Z}}$ to be 3D geometric vectors (positions, velocities, etc) that are steerable by E(3) transformations, and $\boldsymbol{h}$ non-steerable features.

- The function f is E(3)-equivariant, if for any transformation $g \in \mathrm{E}(3)$, $f(g \cdot \vec{\boldsymbol{Z}}, \boldsymbol{h}) = g \cdot f(\vec{\boldsymbol{Z}}, \boldsymbol{h}), \forall \vec{\boldsymbol{Z}} \in \mathbb{R}^{3\times m}, \boldsymbol{h} \in \mathbb{R}^d$.
- Similarly, f is invariant if $f(g \cdot \vec{\boldsymbol{Z}}, \boldsymbol{h}) = f(\vec{\boldsymbol{Z}}, \boldsymbol{h})$.

❑ SubEquivariance

Han et al. (2022a) additionally considers equivariance on the subgroup of $\mathrm{O}(3)$, induced by the external force $\vec{\boldsymbol{g}} \in \mathbb{R}^3$ like gravity, defined as

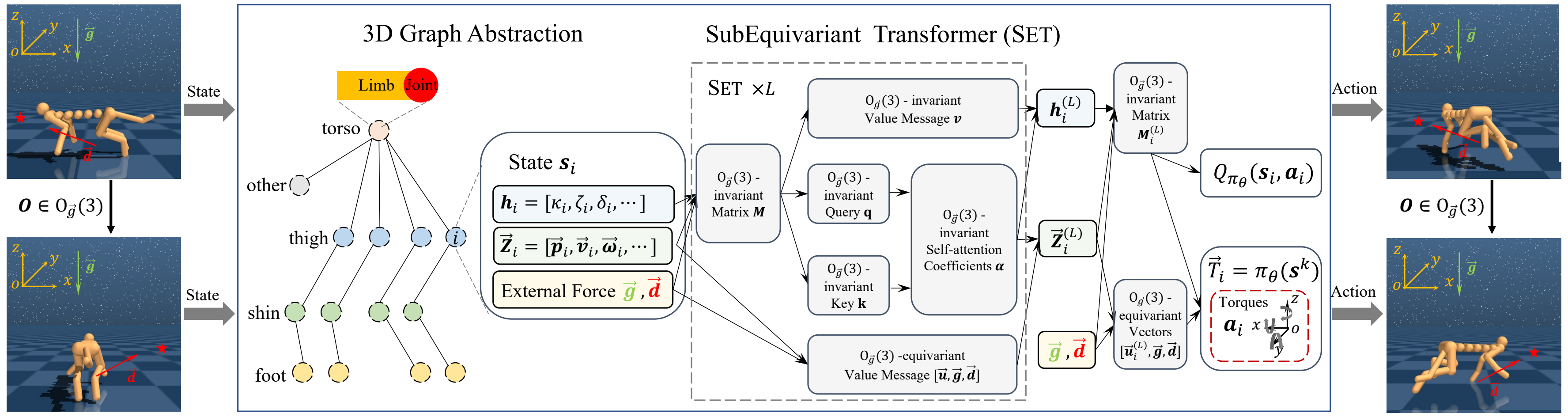
$$\mathrm{O}_{\vec{\boldsymbol{g}}}(3) := \{\boldsymbol{O} \in \mathbb{R}^{3\times 3} \mid \boldsymbol{O}^\top \boldsymbol{O} = \boldsymbol{I}, \boldsymbol{O}\vec{\boldsymbol{g}} = \vec{\boldsymbol{g}}\}$$

Han et al. (2022a) also presented a universally expressive construction of the $\mathrm{O}_{\vec{g}}$ (3)-equivariant functions:

$$f_{\vec{g}}(\vec{\boldsymbol{Z}}, \boldsymbol{h}) = [\vec{\boldsymbol{Z}}, \vec{\boldsymbol{g}}]\boldsymbol{M}_{\vec{g}},$$
$$\text{s.t.} \quad \boldsymbol{M}_{\vec{g}} = \sigma\Big([\vec{\boldsymbol{Z}}, \vec{\boldsymbol{g}}]^\top [\vec{\boldsymbol{Z}}, \vec{\boldsymbol{g}}], \boldsymbol{h}\Big),$$

where $\sigma(\cdot)$ is an Multi-Layer Perceptron (MLP) and $[\vec{\boldsymbol{Z}}, \vec{\boldsymbol{g}}] \in \mathbb{R}^{3\times(m+1)}$ is a stack of $\vec{\boldsymbol{Z}}$ and $\vec{\boldsymbol{g}}$ along the last dimension.

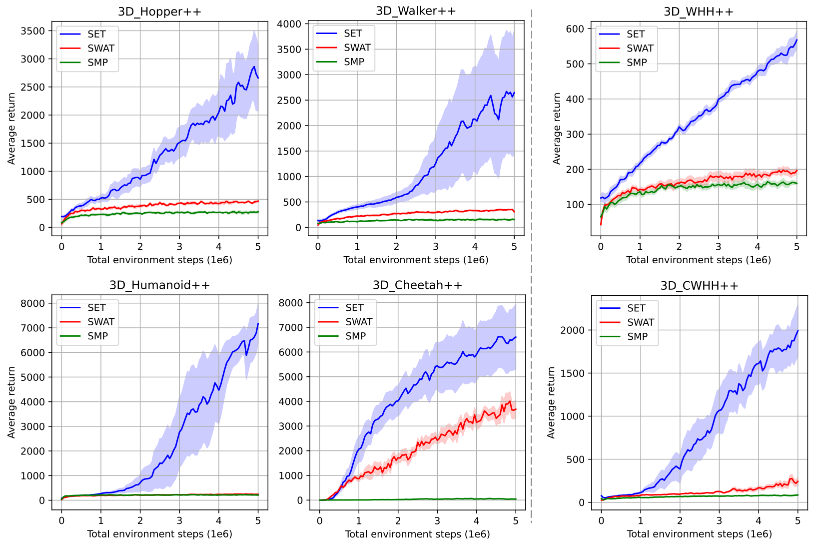
❑ Illustration of the Flowchart of 3D-SGRL.



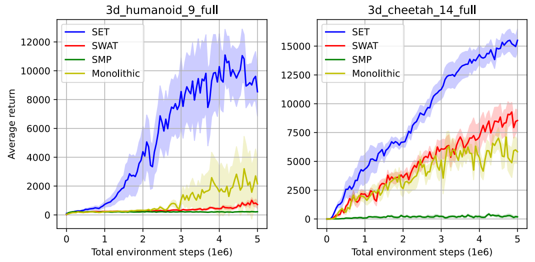
Notation: $\vec{\boldsymbol{Z}}_i \in R^{3\times 6}$ include its position $\vec{\boldsymbol{p}}_i \in \mathbb{R}^3$, positional velocity $\vec{\boldsymbol{v}}_i \in \mathbb{R}^3$, rotational velocity $\vec{\boldsymbol{\omega}}_i \in \mathbb{R}^3$, etc. $\boldsymbol{h}_i \in \mathbb{R}^{13}$ consist of the rotation angles $\kappa_i, \zeta_i, \delta_i$ of joint axes, etc. External forces like gravity $\vec{g} \in \mathbb{R}^3$ and a target direction $\vec{\boldsymbol{d}} \in \mathbb{R}^3$. Also, $\vec{\boldsymbol{u}}_i = \vec{\boldsymbol{Z}}_i \boldsymbol{W}_{\boldsymbol{u}}$, where $\boldsymbol{W}_{\boldsymbol{u}}$ is a matrix.

Experiment

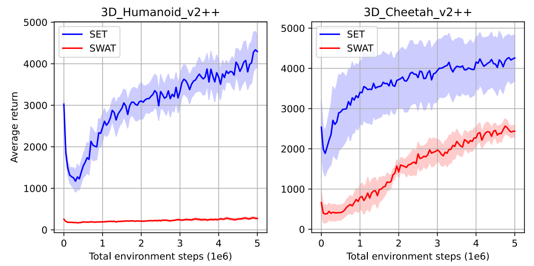
❑ Muti-Task with different Morphologies



❑ Single-Task

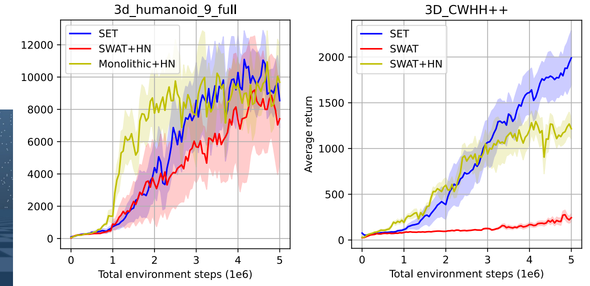


❑ Evaluation on V2-Variants



❑ Comparison with Invariant Methods

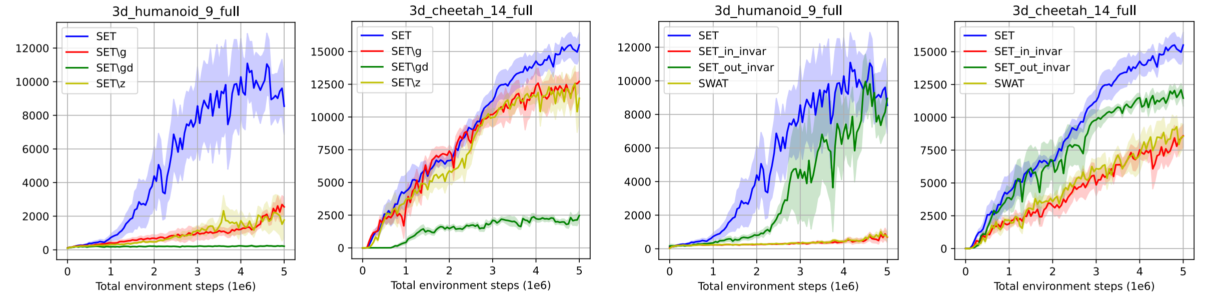
Environment	SET	SWAT+HN
cross-domain (3D_CWHH++)		
3d_walker_3	**206.8** ± 37.4	26.3 ± 72.4
3d_walker_6	**243.7** ± 32.3	156.8 ± 11.1
3d_humanoid_7	**161.9** ± 3.4	130.2 ± 2.1
3d_humanoid_8	**180.0** ± 6.5	152.9 ± 36.8
3d_cheetah_11	**1078.1** ± 722.8	786.5 ± 779.3
3d_cheetah_12	**3038.3** ± 2803.3	2517.3 ± 2113.9



❑ Zero-Shot Generalization

Environment	SET	SWAT	SMP
in-domain (3D_Walker++, 3D_Humanoid++, 3D_Cheetah++)			
3d_walker_3	**276.2** ± 17.4	207.0 ± 52.7	56.8 ± 15.1
3d_walker_6	**431.3** ± 146.2	358.0 ± 58.9	143.4 ± 50.7
3d_humanoid_7	**244.8** ± 7.9	170.3 ± 51.7	190.9 ± 16.2
3d_humanoid_8	**299.6** ± 23.7	141.4 ± 22.1	185.4 ± 9.2
3d_cheetah_11	**4643.9** ± 292.6	1785.3 ± 999.3	2.0 ± 2.9
3d_cheetah_12	**916.0** ± 39.7	744.1 ± 317.1	29.8 ± 10.7
cross-domain (3D_CWHH++)			
3d_walker_3	**206.8** ± 37.4	17.9 ± 13.7	18.0 ± 22.9
3d_walker_6	**243.7** ± 32.3	114.9 ± 40.3	103.9 ± 1.8
3d_humanoid_7	**161.9** ± 3.4	152.0 ± 6.8	124.2 ± 15.7
3d_humanoid_8	**180.0** ± 6.5	156.6 ± 1.7	129.3 ± 0.1
3d_cheetah_11	**1078.1** ± 722.8	4.3 ± 1.6	6.2 ± 0.5
3d_cheetah_12	**3038.3** ± 2803.3	349.7 ± 304.3	6.6 ± 1.2

❑ Ablation



Our Contributions

- ➢We introduce a new morphology-agnostic RL benchmark that extends the widely adopted 2D-Planar setting to 3D-SGRL, permitting significantly larger exploring space of the agents with arbitrary initial location and target direction.
- ➢To learn a policy in this massive search space, we design SET, a novel model that preserves geometric symmetry by construction.

Website

❑ https://alpc91.github.io/SGRL/

❑ https://github.com/alpc91/SGRL

Wechat

Runfa Chen　Jiaqi Han　Wenbing Huang